AMENDMENTS TO THE SPECIFICATION

Please amend the specification as follows:

On page 1, please amend the title as follows:

--METHODS OF CLASSIFYING, DIAGNOSING, STRATIFYING AND TREATING CANCER PATIENTS AND THEIR TUMORS ~~BASAL MARKERS IN BREAST CANCER AND USES THEREOF~~--

On page 4, please amend the paragraph that begins on line 2 as follows:

--The present invention relates to the identification of markers that are useful in classifying tumors, particularly breast tumors. The markers identify a class of tumors whose cells have characteristics of basal cells of normal breast lactation ducts. The markers were identified based on their expression profiles in human breast tumor samples, normal breast tissue, and cell lines as assessed using cDNA microarrays. In particular, the basal cell markers of the present invention were identified based on the similarity of their mRNA expression patterns to the expression patterns of markers previously known to identify breast duct basal cells, e.g., cytokeratin 5 and cytokeratin 17, across a set of breast tumor samples. The basal markers include the three genes known as cadherin 3 or P-cadherin (SEQ ID NO:1; GenBank protein accession number NP_001784 ~~NP_001399~~; GenBank cDNA accession number NM_001793 ~~NM_001408~~), matrix metalloproteinase 14 (SEQ ID NO:2; GenBank protein accession number NP_004986; GenBank cDNA accession number NM_004995); and cadherin EGF LAG seven-pass G-type receptor 2 or EGF-Like Domain, Multiple 2 (SEQ ID NO:3; GenBank protein accession number NP_001399 ~~NP_001784~~; GenBank cDNA accession number NM_001408 ~~NM_001793~~). The invention further provides antibodies that specifically bind to the polypeptides encoded by the basal marker genes identified herein. The antibodies recognize basal cells of normal mammary lactation glands.--

On page 20, please disable the hyperlink on line 7 by amending the paragraph that begins on line 3 as follows:

--Note that Tables 1, 2, and 6 are provided for purposes of presenting the clone identifications and the data that was used to perform hierarchical clustering analysis, and that the format of the tables may not correspond exactly with the format required by software developed for the analysis of the data. Appropriate format will, in general, depend upon the particular computer program. See, for example, the Web site ~~http://~~genome-www.stanford.edu/~sherlock/tutorial.html for discussion of the appropriate format for one particular analysis program.--

On page 30, please disable the hyperlink on line 28 by amending the paragraph that begins on line 1 as follows:

--Several previous studies suggested that expression of basal cell keratins is associated with a poor clinical outcome (Dairkee, S.H., *et al.*, "Monoclonal antibody that predicts early recurrence of breast cancer", *Lancet*, 1:514, 1987; Malzahn, K., *et al.*, "Biological and prognostic significance of stratified epithelial cytokeratins in infiltrating ductal breast carcinomas", *Virchows Archiv*, 433:119-29, 1998). Inventors have confirmed, in a large-scale study, that patients with breast tumors whose cells display characteristics of breast basal cells, e.g., expression of cytokeratin 5 and/or cytokeratin 17, have a poor clinical outcome relative to patients with breast tumors that do not express these markers. However, antibodies to these cytokeratins have been found (by the inventors and by other investigators) to give spotty, focal staining patterns when used to perform immunohistochemistry on breast tumor samples. Thus the utility of cytokeratins 5 and 17 as markers and the utility of antibodies that bind to cytokeratin 5 or 17 for determining whether a tumor is a member of the basal subclass has been limited. The inventors have therefore identified genes whose mRNA expression profiles across a large set of tumor samples correlate with, i.e., are similar to, the expression profiles of the known basal cell markers cytokeratins 5 and 17. These genes include the basal marker genes of the present

invention, i.e., genes that encode cadherin3 or P-cadherin (SEQ ID NO:1; GenBank protein accession number NP_001784 ~~NP_001399~~; GenBank cDNA accession number NM_001793 ~~NM_001408~~), matrix metalloproteinase 14 (SEQ ID NO:2; GenBank protein accession number NP_004986; GenBank cDNA accession number NM_004995); and cadherin EGF LAG seven-pass G-type receptor 2 or EGF-Like Domain, Multiple 2 (SEQ ID NO:3; GenBank protein accession number NP_001399 ~~NP_001784~~; GenBank cDNA accession number NM_001408 ~~NM_001793~~). A portion of the cadherin3 gene was present as I.M.A.G.E. clone 777301 on the cDNA microarray described below. This clone is entry #421 in Appendix H, Table 1. A portion of the matrix metalloproteinase 14 gene was present as I.M.A.G.E. clone 270505 on the cDNA microarray described below. This clone is entry #424 in Appendix H, Table 1. A portion of the cadherin EGF LAG seven-pass G-type receptor 2 gene was present as I.M.A.G.E. clone 175103 on the cDNA microarray described below. This clone is entry #1443 in Appendix H, Table 1. Information about these genes may be found at NCBI's LocusLink (~~http://~~www.ncbi.nlm.nih.gov/LocusLink), among other sources. As described in Examples 10 and 13, the inventors have generated antibodies to the proteins expressed by these genes and shown that the antibodies stain basal cells of normal mammary lactation glands. Thus detection of one or more expression products of these genes may be used to identify tumors that fall within the basal tumor subclass.--

On page 36, please disable the hyperlinks on lines 19 and 28 by amending the paragraph that begins on line 15 as follows:

--Further details of the experimental methods used in the present invention are found in the Examples. Additional information describing methods for fabricating and using microarrays is found in U.S. Patent No. 5,807,522, which is herein incorporated by reference. Instructions for constructing microarray hardware (e.g., arrayers and scanners) using commercially available parts can be found at ~~http://~~cmgm.stanford.edu/pbrown/ and in Cheung, V., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R., and Childs, G., Making and reading microarrays, *Nature Genetics Supplement*, 21:15-19, 1999, which are herein incorporated by reference. Additional

discussions of microarray technology and protocols for preparing samples and performing microrarray experiments are found in, for example, DNA arrays for analysis of gene expression, *Methods Enzymol*, 303:179-205, 1999; Fluorescence-based expression monitoring using microarrays, *Methods Enzymol*, 306: 3-18, 1999; and M. Schena (ed.), DNA Microarrays: A Practical Approach, Oxford University Press, Oxford, UK, 1999. Descriptions of how to use an arrayer and the associated software are found at ~~http://~~cmgm.stanford.edu/pbrown/mguide/arrayerHTML/ArrayerDocs.html, which is herein incorporated by reference.--

On page 51, please disable the hyperlink on line 24 by amending the paragraph that begins on line 11 as follows:

--Various other techniques for detecting the basal marker polypeptides identified herein are within the scope of the invention. For example, a basal marker polypeptide may be detected using an assay for a biochemical activity of the polypeptide, e.g., an enzymatic activity. This type of assay may be especially convenient for tests on samples such as blood or other body fluids. Such an approach may be particularly attractive in the case of matrix metalloproteinase 14. As described above, matrix metalloproteinases are involved in cleavage of various proteins in the extracellular matrix. The cleavage specificity of this protein may readily be determined, and an appropriate substrate prepared. (See, e.g., Turk, B., *et al.*, "Determination of protease cleavage site motifs using mixture-based oriented peptide libraries", *Nature Biotechnology*, 19(7): 661-667, 2001, which discusses cleavage site motifs for various metalloproteases including MMP14, referred to as MT1-MMP therein.) Cleavage of this substrate may then be detected. In certain embodiments of the invention the substrate includes a fluorescent moiety for convenient detection. The invention contemplates use of fluorescent resonance energy transfer (FRET) assays to detect matrix metalloproteinase 14 (see ~~http://~~www.aurorabio.com).--

On page 65, please disable the hyperlink on line 11 by amending the paragraph that begins on line 8 as follows:

--The human cDNA clones used in this study were obtained from Research Genetics (Huntsville AB, USA) as bacterial colonies in 96-well microtiter plates. The clones were chosen from a set of 15,000 cDNA clones that corresponded to the Research Genetics Human Gene Filters sets GF200-202 (~~http://~~www.resgen.com/). These clones form part of a set of clones assembled by the I.M.A.G.E. consortium (Lennon, G.G., Auffray, C., Polymeropoulos, M., Soares, M.B. The I.M.A.G.E. Consortium: An Integrated Molecular Analysis of Genomes and their Expression. Genomics 33:151-152,1996) and are identified by I.M.A.G.E. clone ID numbers. All clones printed on these arrays were sequence validated as part of a product offered at Research Genetics, Inc. We estimate that greater than 97% of the clones on the array are correctly identified.--

On page 65, please disable the hyperlink on line 18 by amending the paragraph that begins on line 17 as follows:

--A detailed protocol for the production of the cDNA microarrays used in this study is available at ~~http://~~cmgm.stanford.edu/pbrown/protocols.html and is reproduced below with insubstantial changes. As described below, the protocol includes steps of (1) cleaning the glass slides onto which the DNAs (e.g., products of PCR reactions) are to be spotted; (2) spotting the DNAs onto the glass slides with an arrayer; (3) Post processing to prepare arrays containing spotted DNAs for hybridization. All procedures are done at room temperature and with double distilled water unless otherwise stated. Unless otherwise stated, in this Example and the following Examples, reagents are prepared according to protocols available in Maniatis, T., Sambrook, J. and Fritsch, E., *Molecular Cloning: A Laboratory Manual* (3 Volume Set), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989. --

On page 69, please disable the hyperlink on line 12 by amending the paragraph that begins on line 9 as follows:

-- Following their excision, breast tumor samples were rapidly frozen in liquid N2 and then stored at -80 C until use. mRNA was isolated from breast tumors and normal breast tissue using the Trizol Reagent (Gibco-BRL) and Invitrogen FastTrack 2.0 Kit (all Stanford samples, and see ~~http://~~genome-www.stanford.edu/sbcmp/web.shtml for the detailed protocol) or using the Trizol Reagent followed by Dynal bead separation for the mRNA purification step (all Norway tissue samples). Briefly, frozen tumor samples were cut into small pieces and immediately placed into 12 ml of Trizol Reagent. Each tumor sample in Trizol was homogenized using a PowerGen 125 Tissue Homogenizer (Fisher Scientific), and total RNA was isolated according to the Trizol reagent manufacturer's protocol. Tumor mRNA was isolated according to the manufacturer's protocols using the FastTrack 2.0 Kit (Invitrogen) or Dynal beads. --

On page 70, please disable the hyperlink on line 2 by amending the paragraph that begins on page 69, line 25 as follows:

--A single pathologist (applicant Matt van de Rijn) reviewed hematoxylin and eosin (H&E) sections of each tumor, including all before and after pairs, and made a histological evaluation of each while blinded to the source. Tumors were graded using a modified version of the Bloom-Richardson method (Robbins, P., et al., *Hum Pathol*, 26, 873-879, 1995). These data are displayed in Appendix H, Table 4. Representative H&E sections of each tumor are posted on Applicants' website at ~~http://~~genome-www.stanford.edu/molecularportraits/.--

On page 72, please disable the hyperlink on line 15 by amending the paragraph that begins on line 13 as follows:

--22. During Step 21, prepare the necessary number of hybridization chambers (Custom made by Die-Tech, San Jose, CA (see "Drawings for custom parts at ~~http://~~cmgm.stanford.edu/pbrown/mguide/HybChamber.pdf") or purchased at Corning Costar, Acton, MA (CTM™ Hybridization Chamber, #2551), get 22mm X 22mm coverslips ready, and get arrays ready.--

On page 73, please disable the hyperlinks on lines 12, 14 and 16 by amending the paragraph that begins on line 8 as follows:

--The cDNA microarrays were scanned with either a General Scanning (Watertown, MA) ScanArray 3000 at 20 microns resolution, or with a prototype Axon Instruments (Foster City, CA) GenePix Scanner at 10 micron resolution. The output files, which were TIFF images, were then analyzed using the program ScanAlyze (M. Eisen; available at ~~http://~~www.microarrays.org/software). Fluorescent ratios and quantitative data on spot quality (see ScanAlyze manual) were stored in a prototype of the AMAD database (M. Eisen; available at ~~http://~~www.microarrays.org/software). Areas of the array with obvious blemishes were manually flagged and excluded from subsequent analyses. The primary data tables can be downloaded at ~~http://~~genome-www.stanford.edu/molecularportraits/, in text/tab delimited format after obtaining a password.--

On pages 73-74, please disable the hyperlinks on page 73, line 29 and page 74, line 2 by amending the paragraph that begins on page 73, line 28 as follows:

--Average-linkage hierarchical clustering, as implemented in the program Cluster (M. Eisen; ~~http://~~www.microarrays.org/software), was applied separately to both the genes and arrays. The results were analyzed, and images generated, using TreeView (M. Eisen; ~~http://~~www.microarrays.org/software).--

On page 100, please disable the hyperlink on line 25 by amending the paragraph that begins on line 20 as follows:

--The following cell lines were used: 184B5, MCF7, OVCAR3, UACC62, HepG2, Colo205, UACC62, JURKAT, N-TERA2, MOLT4, Sw872. These cell lines are well known in the art. Descriptions of these cell lines are provided in Table 3, in Perou, et al., Molecular

portraits of human breast tumours, *Nature*, 406(6797):747-52, 2000, in Ross, D. T. et al. Systematic Variation in Gene Expression Patterns in Human Cancer Cell Lines. Nature Genetics, 24(3):227-35, 2000, and at the American Type Culture Collection Web site: ~~http://~~www.atcc.org. Cell lines were maintained under standard growth conditions and in standard tissue culture media as appropriate for the particular cell line. Cells were collected according to standard techniques (e.g., trypsinization in the case of adherent cells), and the resulting cell suspension was prepared as follows:--